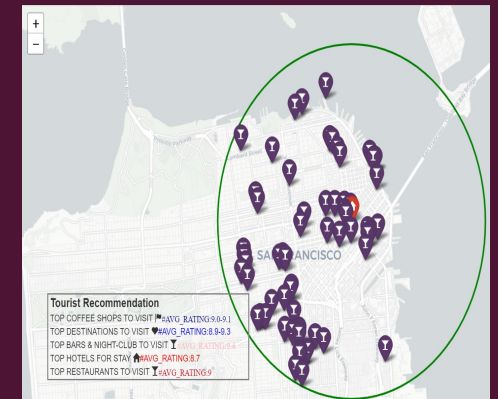
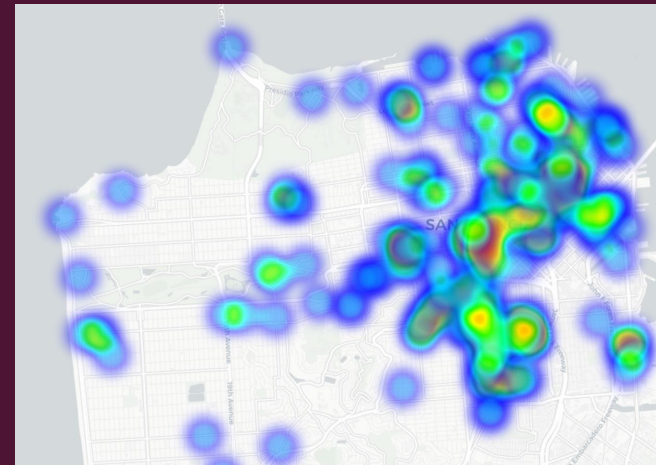
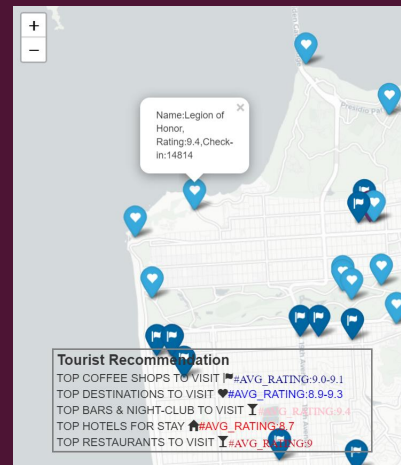
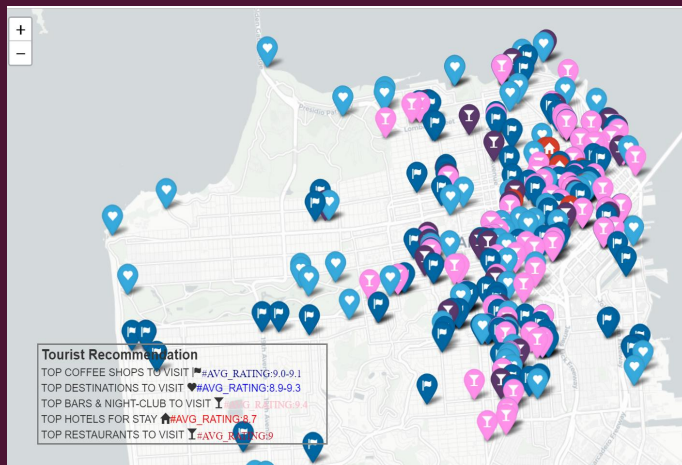


VISIT SAN FRANCISCO: WHAT TO SEE, WHERE TO STAY & WHERE TO EAT.

SAN FRANCISCO'S VENUES RECOMMENDATION MAPS



RECOMMENDING VENUES TO TOURIST OF SAN FRANCISCO

- A person wanted to visit San Francisco we would help him by recommending places to stay(Hotels), visiting places and best place to eat.
- Such places is recognized by a criteria : Most Visits, Cost, Number of Likes and Rating (given by visitors)
- Recommendation would help visitors in many ways.
 - Save their precious time in search of venues.
 - Can budget their visit accurately.
 - Plan their schedule easily.
 - Able to navigate easily by the help of the MAP.
- Map also identifies areas which are close to staying locations

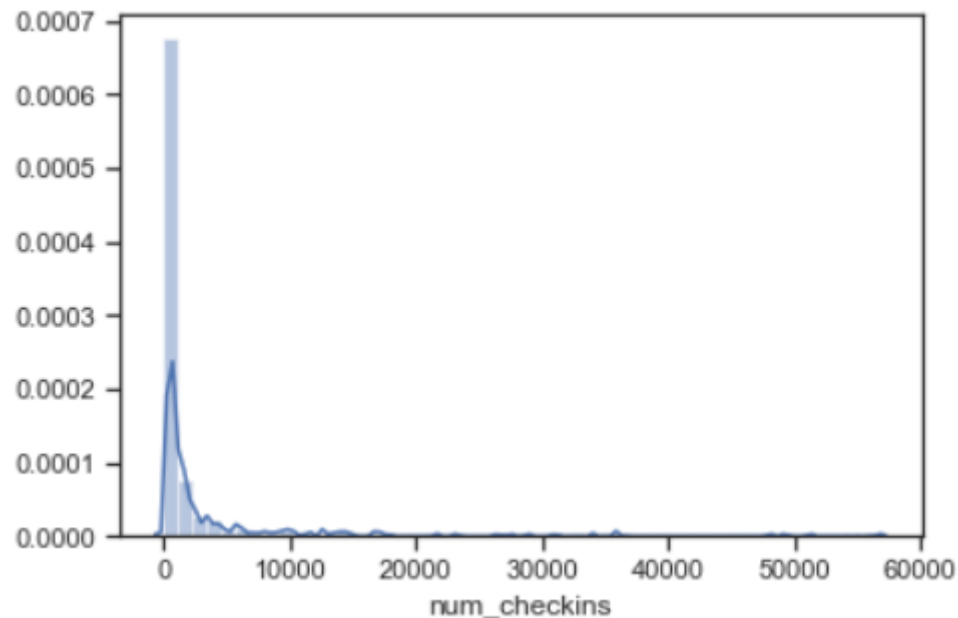
DATA GATHERING & DATA ACQUISITION

- Venue id, location(lat, log), categories and name from FOURSQUARE, data scraped from [FOURSQUAR](#)
- Getting San Francisco Venue detail data Foursquare API from [FOURSQUAR](#) which comprised of name, categories, check-ins count, likes count ,location(lat, log), pricing, rating, venue URL and url_foursquare.

	id	name	categories	lat	long	num_checkins	num_likes	price	rating	url_venue	url_foursquare
0	4ec6e39799115348d857eef7	Pioneer Monument (James Lick Monument)	Art Gallery	37.779800	-122.416135	43	0.0	NaN	NaN	http://www.sfartscommission.org	http://4sq.com/sCDvwW
1	4ee2341cf9ab478e5935d27d	Calibre One	Office	37.783210	-122.393368	10	0.0	NaN	NaN	NaN	http://4sq.com/t4zgoG
2	4e9126889a52db7dba8dc0a0	The Red Room	Nightlife Spot	37.734963	-122.479835	7	0.0	NaN	NaN	NaN	http://4sq.com/qO5NSG
3	53503f72498ef2dd1139298b	Strip, Cuddle Up And Nap	Residential Building (Apartment / Condo)	37.712424	-122.430949	2	0.0	NaN	NaN	NaN	http://4sq.com/1lcZpnN
4	4a1f1787f964a520057c1fe3	Madusalon	Salon / Barbershop	37.772272	-122.437238	529	17.0	NaN	7.4	http://www.madusalon.com	http://4sq.com/5mEGxl

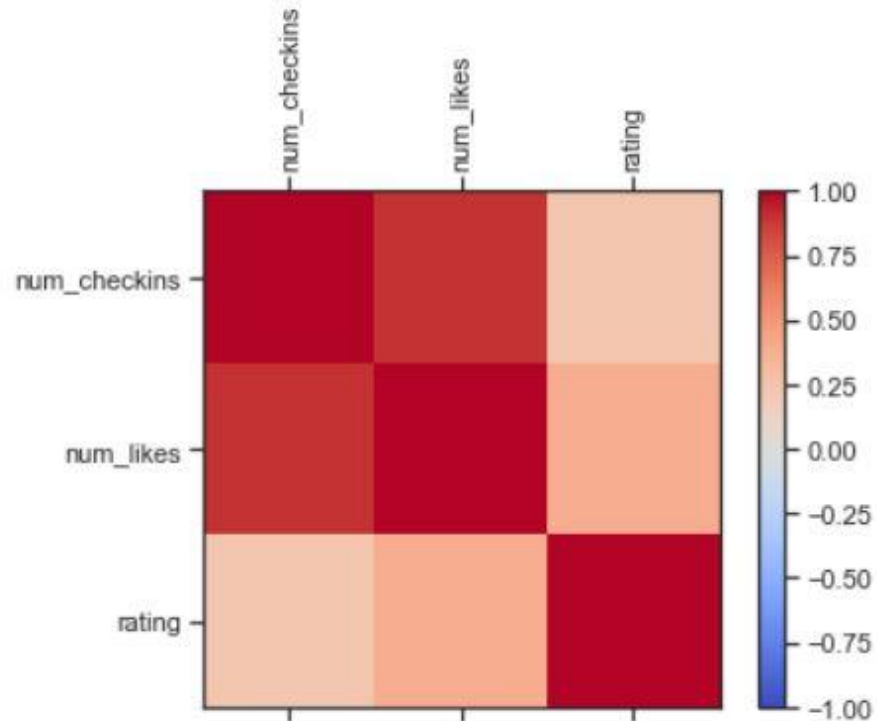
DATA ACQUISITION

- In total, 44,425 rows and 11 features in the raw dataset.
- Null values and duplicated data were dropped.
- Segregating relevant data like staying places, restaurants, coffee shops, and visiting venues.
- After checking 'check-in' data comes to know that most of them values are below 1000, so filtering data on number of visitor: minimum above 1000 and rating above 6.9 out of 10.



CHECKING CORRELATION MEASURE

- Finding showed no or very little correlation between different features.



```
[27]: pg.corr(x=df['num_likes'], y=df['num_checkins'])
```

```
[27]:
```

	n	r	CI95%	r2	adj_r2	p-val	BF10	power
pearson	67	0.902	[0.84, 0.94]	0.813	0.807	2.256251e-25	1.454e+22	1.0

```
[28]: pg.corr(x=df['num_likes'], y=df['rating'])
```

```
[28]:
```

	n	r	CI95%	r2	adj_r2	p-val	BF10	power
pearson	67	0.394	[0.17, 0.58]	0.155	0.129	0.000967	31.409	0.919

```
[29]: pg.corr(x=df['num_checkins'], y=df['rating'])
```

```
[29]:
```

	n	r	CI95%	r2	adj_r2	p-val	BF10	power
pearson	67	0.231	[-0.01, 0.45]	0.053	0.024	0.059916	0.862	0.474

PRE-PROCESSING BEFORE USING K-MEANS ALGORITHM

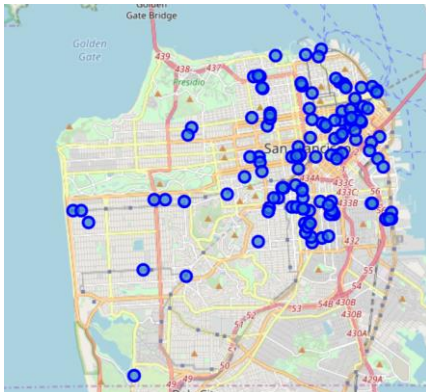
- We wanted to cluster each category on the basis of available features.
- Before using K-mean clustering removing values have no part in making cluster like name, id, categories, lat, long, price, url_venue, url_foursquare.
- Inverting that price(price value from 1 (least pricey) - 4 (most pricey)) information so that overall increase trend show same as rating, which would ultimately help algorithm to cluster data in increase value. i.e. when value is '4' replace it with '1'
- Used Min-Max for scaling.

CLUSTERING USING K-MEANS ALGORITHM

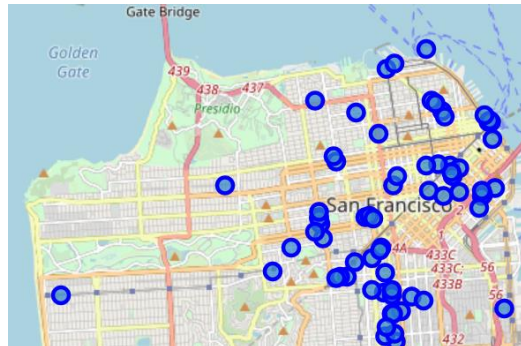
Top venues clustered.

Following pictures showing cluster report and their locations.

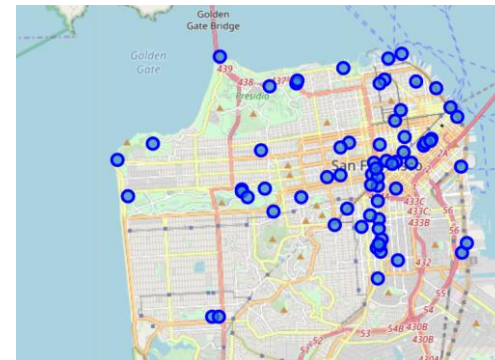
cluster_Small_shops_Report_s					
	Cluster	Mean_Rating	Checkins	Mean_likes	Mean_price
1	1.0	9.160000	62938.0	1487.320000	1.520000
3	3.0	9.052727	28545.0	475.727273	2.118182
4	4.0	8.861475	23003.0	235.844262	1.000000
6	6.0	8.405797	24371.0	205.050725	2.159420
2	2.0	8.034965	27368.0	99.685315	1.000000
0	0.0	7.520000	22939.0	116.234783	2.121739
5	5.0	7.312879	21360.0	57.924242	1.000000



cluster_Gourmet_life_Report_s					
	Cluster	Mean_Rating	Checkins	Mean_likes	Mean_price
4	4.0	9.081429	56439.0	903.400000	1.942857
3	3.0	8.843796	20519.0	277.255474	3.197080
2	2.0	8.576224	24371.0	205.489510	2.000000
0	0.0	8.558929	16073.0	172.080357	1.000000
6	6.0	7.809302	17321.0	138.593023	3.127907
1	1.0	7.543077	21157.0	97.534615	2.000000
5	5.0	7.468182	12721.0	54.032468	1.000000



cluster_Night_life_Report_s					
	Cluster	Mean_Rating	Checkins	Mean_likes	Mean_price
2	2.0	9.082353	56439.0	972.764706	2.029412
6	6.0	8.723529	20519.0	303.117647	3.137255
0	0.0	8.562295	24427.0	233.336066	2.000000
4	4.0	8.426923	20853.0	195.961538	1.000000
3	3.0	7.651351	17321.0	144.702703	3.054054
1	1.0	7.511765	27402.0	139.226891	2.000000
5	5.0	7.475439	27988.0	97.368421	1.000000

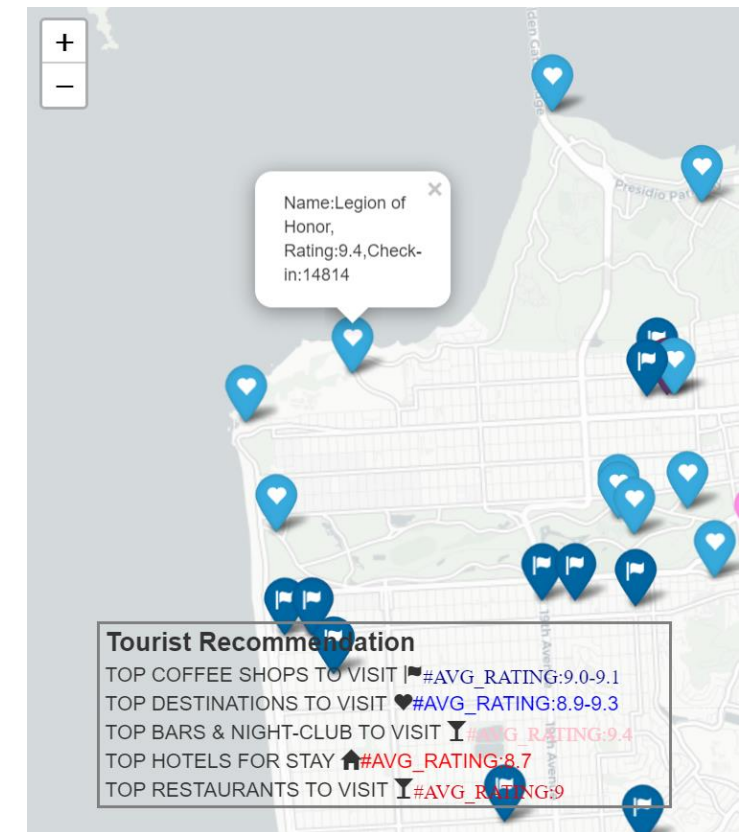
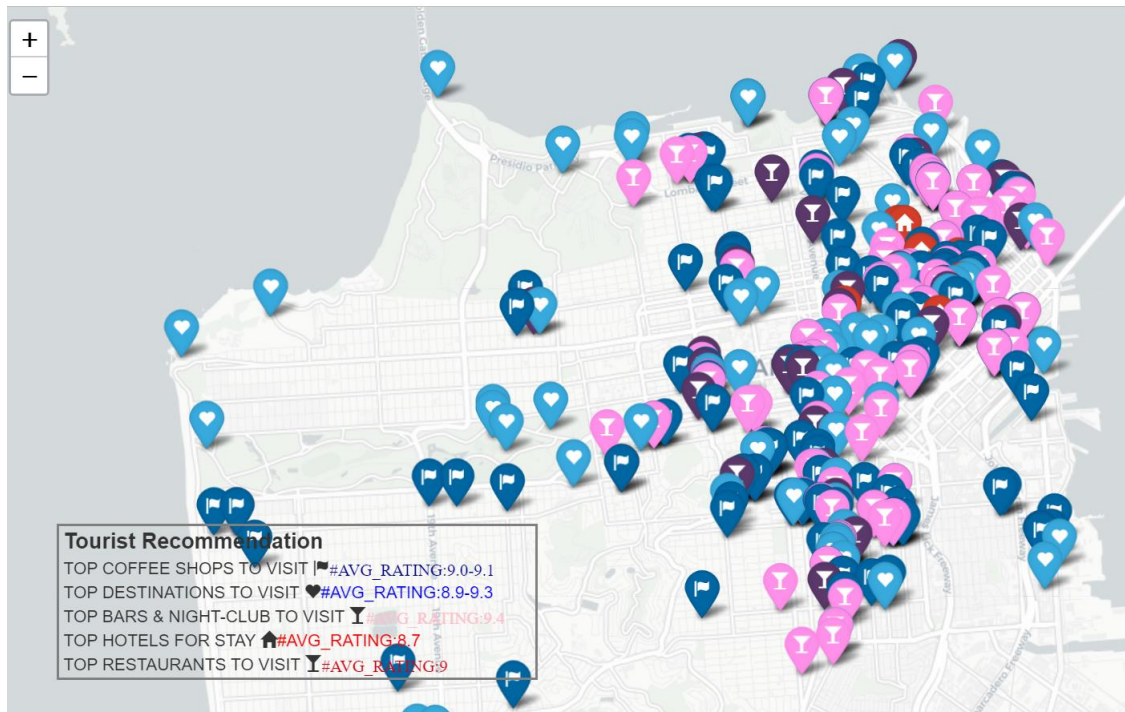


MAKING TOURIST RECOMMENDATION MAP

Following is Tourist recommendations MAP of San Francisco giving information.

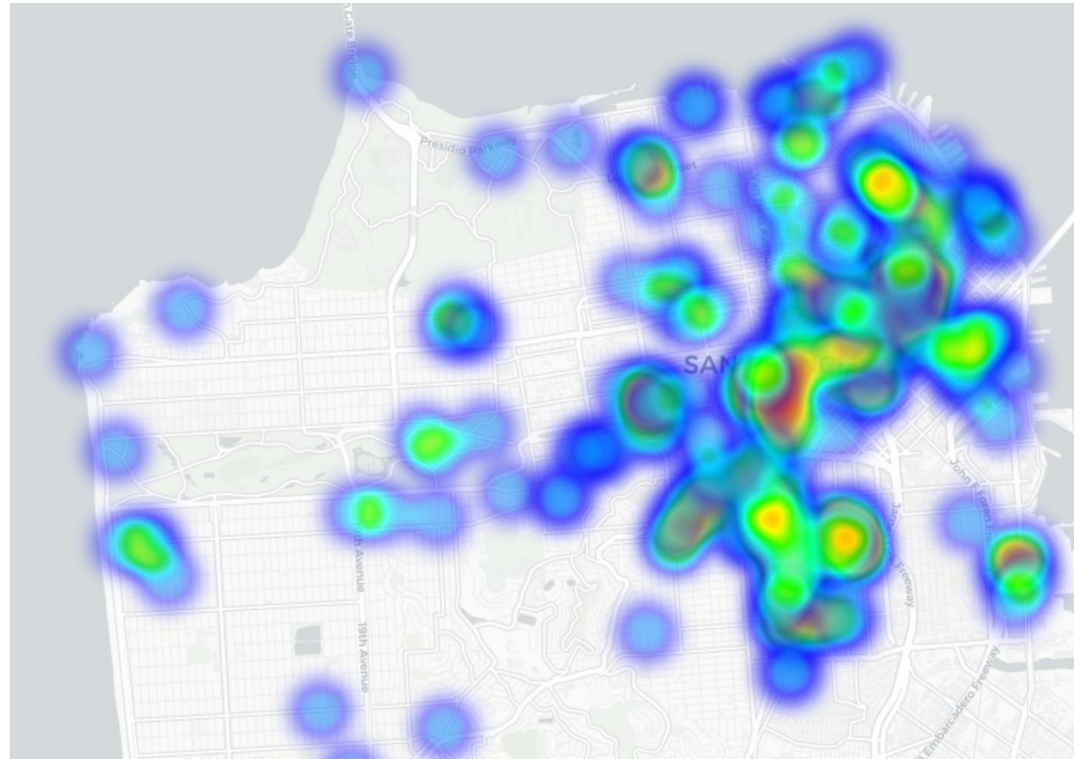
- Hotels, restaurants,
- Tourist attractive places.
- Nigh-time activities i.e. bars, night clubs
- Coffee shops

Along with Rating & Number Check-in information.



MAKING TOURIST RECOMMENDATION MAP

Making heat map which will help to see density of highly recommended venues. Visitor can select his/her staying accordingly



FIND OPTIMAL LOCATIONS NEAR VISITORS STAYING PLACE

- Made a maps which will help visitors to chose specific venues within 4 km of staying location.
- First calculated distance from a hotels to all restaurants.

[162]:

	Name	lat	long	Rating	Checkins	likes	Price	DistanceFromRegis
0	Southern Pacific Brewing	37.760161	-122.414028	8.7	33749.0	1263.0	2.0	3.075243
1	Gott's Roadside	37.796026	-122.394165	8.7	25185.0	705.0	2.0	1.289966
2	Suppenküche	37.776346	-122.426190	9.1	16838.0	757.0	2.0	2.423324
3	15 Romolo	37.798102	-122.406378	8.8	19116.0	599.0	2.0	1.418258
4	Buena Vista Cafe	37.806731	-122.420775	9.2	24295.0	979.0	2.0	2.867338

- Filtering restaurants whose distance is less than 4km.

FIND OPTIMAL LOCATIONS NEAR VISITORS STAYING PLACE

- Making map which marking Hotel and surrounding restaurants within 4 km circle.

